



MCESA

Maricopa County Education Service Agency

Assessment Development Webcast Script

Welcome. The purpose of this webcast is to familiarize you with the contexts and processes that contributed to the development of assessments used in your district for teachers of art, music, PE, band, choir, theater and dance. These assessments are the result of a county-wide collaborative project coordinated by MCESA.

MCESA is the Maricopa County Education Service Agency. We are the office of the Maricopa County School Superintendent, Dr. Don Covey. MCESA has three areas of focus: Executive Leadership, which oversees school board elections and home schooling; Economic Management, which assists districts with financial management; and Educational Innovations, which services school districts and charter schools through grant funded initiatives and a regional training center. One of our many functions is to support school districts in the county with the implementation of state-wide initiatives and legislation.

In particular, the state of Arizona's legislature passed Arizona Revised Statute 15-203. This statute says...

"The State Board of Education shall on or before December 15, 2011 adopt and maintain a model framework for a teacher and principal evaluation instrument that includes quantitative data on student academic progress that accounts for between thirty-three percent and fifty percent of the evaluation outcomes and best practices for professional development and evaluator training. School districts and charter schools shall use an instrument that meets the data requirements established by the State Board of Education to annually evaluate individual teachers and principals beginning in school year 2012 – 2013."

The *Arizona Framework for Measuring Educator Effectiveness* is a 37-page document describing the details and expectations of this statute. In this document teachers are categorized as either Group A teachers, those with valid and reliable classroom level student achievement data aligned to the state standards or as Group B, those with limited or no available valid and reliable classroom level data aligned to state standards.

In a survey conducted in the fall of 2011, MCESA found that 74% of Maricopa County districts requested assistance with the development of assessments to comply with ARS 15-203, and thus began the project to develop assessments to support teachers in Group B. Press pause here to discuss the first two questions on the handout that accompanies this webcast.

- Who is MCESA?
- Why have they facilitated the development of assessments for special area courses?



MCESA

Maricopa County Education Service Agency

The very first step was to create the Cross District Assessment Advisory Council. This council included administrators for research and assessment from across county school districts and nationally known consultants in the field of assessment. This council helped to determine the design and format of the project and the assessments.

When designing assessments, it is necessary to consider the whole framework of balanced assessment as a way to guide decisions. Any balanced assessment plan recognizes the value of a variety of formats including selected response items such as multiple choice or true-false items, constructed response items such as essays and short answer and performance-based assessments. All of these formats were considered when initially planning the design of the assessments to be created. Multiple choice tests were chosen for the first stage of the development process for several reasons. They are cost effective to create. They can be administered in a secure on-line format. They require minimal training for test administration. They align to the format of present national and high stakes tests. They are easy to score, and they provide a good base for critical thinking. Also, the cost of training and implementing constructed response and performance-based assessments were prohibitive for many districts in our current economic environment.

Given all of these strong reasons, it was decided to begin our collective assessment work with multiple choice tests of approximately 45 questions to be administered in one class period with the intention to investigate the use of performance-based assessments in the future. Furthermore, MCESA partnered with West ED, an assessment development company, to design and employ a research-based process to create assessments that are highly reliable and valid to ensure assessments linked to teacher effectiveness are of the highest quality.

Reliability means the degree to which a test produces similar scores each time it is used. Validity means the extent to which a test measures what it is intended to measure. To create valid and reliable test items, MCESA and WestED employed three key strategies in the development process. 1.) Align assessments to grade and content standards and Depth of Knowledge levels. 2.) Use professional editors to monitor style and formatting. 3.) Follow research-based guidelines for writing stems, answers and distractors in collaboration with special area teachers. Press pause here to discuss the third question on the handout that accompanies this webcast.

- What are two things that support the quality of the assessments?

Let's look more closely at the development process for the MCESA post-assessments. West Ed provided training and facilitated the work of more than 200 teachers and administrators representing 40 Maricopa County school districts. These teaching professionals worked to create 11 different assessments in these areas:

- Art for 3rd grade, 8th grade and high school
- PE for 3rd grade, 8th grade and high school
- Music for 3rd grade
- Beginning band, beginning choir, high school theater, and high school dance



MCESA

Maricopa County Education Service Agency

The first step was prioritizing standards. This required teams of teachers to sort and organize the Arizona State Standards for their content area. In this step, teams had to ...

- Agree on the most important things students should know and be able to do,
- Agree on the most important things to assess, and
- Determine the relative instructional emphasis, which means to prioritize the standards.

Because the standards for art, dance, PE, band and choir are grouped into grade bands or proficiency levels, the teams had to spend considerable time discussing how the standards look at the different grades within the band. Once the standards were prioritized, teachers were able to use the prioritization to write an assessment blueprint. This document shows the priority of the strands and concepts of the standards through a percentage. For example this blueprint shows that 56% of the items, which is 25 out of 45 items, on the 3rd grade music assessment will be testing Strand 1: Create. Further down on the blueprint you can see that all 25 of those questions will be connected to Concept 5 of Strand 1: Reading and notating music. The other concepts have a zero listed, because the test will not contain any items aligned to those concepts. An end-of-course or end-of-grade assessment would be unmanageable if it tested every performance objective. That is why teacher prioritization of the standards informed the blueprint. Also, not every performance objective in the special areas can be assessed using multiple choice. Thus, it is important to have a balanced assessment plan.

The next step was creating item specifications. Item specifications serve as a bridge between the standards and the actual test items.

- They provide interpretations and clarification of the standards;
- They define the assessable content for each standard; and
- They serve as a collection point for decision rules related to the standards.

To write an item specification, a team may have to either narrow or expand the content of the standard. Furthermore, they specify the essence of the level of cognition for each standard. For example this item spec from 8th grade art shows the standard, the specifications, and the Depth of Knowledge, abbreviated as DOK.

- DOK 1 means the standard focuses on facts and simple recall.
- DOK 2 means the standard focuses on reasoning, such as comparing or applying factual knowledge.
- DOK 3 means the standard focuses on complex thinking and the application of skill.

After the item specifications were created, work began on writing the actual test items. Each item contains a stem and four answer choices. One of the choices is the correct answer, and the other three choices are called distractors. Each item is also coded with a DOK level. Writing items at the DOK 2 level was encouraged. The vocabulary of each item was monitored to be grade-level appropriate and aligned to the content standards. For the MCESA assessments, 165 items were created for each content area. It is necessary to make so many items, because not all of them pass the reviews and field testing process. The goal of the item-writing process was to finish with two high-quality forms of each assessment.

Once the items were written, they were edited by the West Ed content and assessment editors, using the Arizona Style Guide. The items were also formatted to be loaded in the electronic delivery system. Copyright permissions for any texts or graphics were also obtained in this stage.



MCESA

Maricopa County Education Service Agency

After editing, the items were put through a bias review. Approximately 25 community members representing the diversity of Maricopa County examined each item to evaluate whether or not any item created an unfair advantage or disadvantage for any subgroup of students. Items were examined for stereotypes, insensitivity, gender imbalance, and offensiveness. A content review immediately followed the bias review. Teams of teachers who are experts in their content area reviewed each item for clarity, alignment to the standard, DOK level and answer choices. They ensured that the correct answer was listed and the other choices were plausible and not too obvious. After the reviews, the test items were revised and edited once more, and actual test forms were created for field tests.

In May 2012, we field tested the post-assessments for 3rd grade art, music and PE, 8th grade art, 8th grade PE and choir. Twelve districts in Maricopa County participated including the REIL grant partners: Alhambra, Gila Bend, Nadaburg, Tolleson, and Isaac as well as Cave Creek, Dysart, Laveen, Pendergast, Phoenix Elementary, Wilson, and Saddle Mountain. Approximately 12,000 students in over 437 classes participated. Psychometric analysis of the field test data provided reliability coefficients for each test form. Basic statistical practice expects test scores to have a reliability coefficient of .80 or higher. You can see that the results from field testing meet that expectation, indicating these tests have high reliability.

The last step in the assessment development process is to create the operational forms of tests. Any items that did not perform well in the field test were removed and two forms of each test were created for operational use using the quality items from the field-tested versions of the assessment. The following summative or post-assessments will be available for operational use in the 2012-13 school year: 3rd grade art, music and PE and 8th grade art and PE.

In May of 2013, MCESA, ATI and WestEd will field test the following assessments: beginning band and choir, as well as high school art, PE, theater and dance.

Considering the rigorous and research-based process used to develop these post-assessments, you can feel confident that these assessments are a quality product fulfilling one component of a balanced assessment plan. Press pause here to discuss the questions 4 and 5 on the handout that accompanies this webcast.

- What part of the assessment development process contributes to valid content?
- What part of the assessment development process contributes to fair and reliable items?

In addition to the post assessments described here, 11 corresponding pre-assessments were developed in the summer of 2012. These pre-assessments will be field tested in October of 2012. To learn more about assessment development, consider consulting the reference materials provided on the handout.

Also, MCESA will be offering a one-day workshop for teachers or leaders of art, music and PE on October 2nd or 3rd, 2012. The purpose of the workshop is to closely examine the actual item specifications and blueprints and learn how to use them as a curriculum and instruction support for teaching of the content standards. Contact your district administrator if you are interested in joining the workshop on one of these days. To help you process all of the information provided here, answer the remaining questions on your handout.

Thank you for your interest in assessment development.